Sebastian Scholz:

Learning and Technology Adoptions

Munich Discussion Paper No. 2010-5

Department of Economics
University of Munich

Volkswirtschaftliche Fakultät
Ludwig-Maximilians-Universität München

Learning and Technology Adoptions*

Sebastian Scholz†
MGSE, University of Munich

January 13, 2010

Abstract

This essay studies the optimal timing for a firm to adopt a new process innovation in the presence of learning. A policy that has been implemented by governments throughout the world to reduce the cost level of infant industries with positive externalities, is to either subsidize the research of these technologies or their distribution. This model demonstrates how government interventions can affect the optimal timing for adoption of a new technology. Furthermore this essay makes predictions on how the effects change, when the total quantity that can be produced is fixed; the installations of wind powered energy plants exemplify this point. Depending on whether producer rents, consumer rents or early implementation are more important to the government, the model offers the appropriate tools to attain its objective.

Keywords: Learning, Process Innovation, Optimal Control, Infant industry
JEL Classifications: C61, D42, H23, O12

1 Introduction

Typical infant industries are characterized by cost reductions through learning in the production process, and continuous new technology adoptions. Mature industries are often characterized by numerous technology generations, while learning takes place at the same time.[1] Market players try to find new technologies that are more sustainable, efficient and safer, however, at the same

*This paper was written with the support of the Deutsche Forschungsgemeinschaft (SFB/TR 15). For helpful comments and discussions, I thank Ray Rees, Monika Schnitzer, Sven Rady and seminar/ conference participants at the University of Munich, Intertic (Milan, 2008), EARIE (Toulouse, 2008) and INFORMS (Washington DC, 2008). All remaining errors are mine.

†Munich Graduate School of Economics, Kaulbachstr. 45, 80539 Munich, Germany; Email: Sebastian.Scholz@campus.lmu.de

[1] Currently produced nuclear power plants for example use the 3rd technology generation, the 4th generation will be deployed some time around 2030.

time they are improving existing technologies. Thus it is important to connect both: experience and innovations in a single model. In this dynamic framework a firm can adopt innovation breakthroughs from its research department. In addition, the firm decides upon a pricing rule for each point in time. It is assumed that experience spills over to the next technology generation after an innovation breakthrough has been adopted. The empirical literature till present, has concentrated on learning models, in which technology spillovers were absent (Irwin and Klenow, 1994). Jamasb (2007) is an exception: in his purely econometric analysis, he estimates learning by doing and research rates for a range of energy technologies in different stages of technical progress. He separates the cost reduction effect caused by learning and research, expressed by cumulative sales and patents. Unfortunately, it is difficult to obtain data on costs, which makes the study rely on very few data points.

This model shall be the theoretical foundation of applied work, in which firms can use the experience, they have accumulated thus far, for the next technology generation. In macroeconomics, there are studies, where the experience gained from learning, is passed on from one generation to the next. Examples are Young (1993) and Parente (1994). Young makes clear that innovations occur in markets that are large. In this setup, production costs do not decrease with new technology adoptions, but rather through learning. In Parente's model, learning and technology adoptions occur both after the product introduction. A firm faces a trade-off between learning at a decreasing rate or switching to a new technology, which is costly as not all expertise can be transferred. In return, the learning curve becomes steeper. From a microeconomic perspective, both models face one problem in particular: learning occurs only through time and not through cumulative production or "by doing". Therefore the strategic pricing behavior of firms can not be analyzed. This paper introduces a model, where firms simultaneously choose a research budget and the optimal production quantity, exploiting the learning effect optimally. The production and the time of technology adoptions are control variables of the firm.

This research has two main objectives: firstly, to describe the market equilibrium of a setup that accounts for innovations and learning; and secondly, to show the effects of subsidies on the market equilibrium. The second objective is based on the observation that products produced by learning industries have often rendered positive externalities in the past; renewable energy technologies can be cited as examples. The production cost per unit of electricity has been reduced significantly for technologies that are powered by wind, sunlight and biomass. The positive externality is the deduction of the carbon dioxide level in the atmosphere, because electricity from renewable energy is a perfect substitute to conventionally generated electricity.[2] It is illustrated that sales and innovation subsidies have the same effect on the innovation date and prices, if and only if the innovation date and total cumulative production quantity are

[2] Another example is the aerospace technology, which was mainly developed for military purposes during the 1930s and 40s. This was a stepping stone for the development of commercially used airplanes, which has enabled societies to travel and trade at an increased pace. The learning effect in this industry was described by Wright (1936).

endogenous. Effects differ significantly, when the total quantity that can be produced in a market is fixed. An example for such products, whose costs are affected by learning are wind power plants. In Germany the installation of onshore wind power plants reached its peak in 2002, with an installed capacity of over 3000 MW. The installed capacity in 2009 was estimated to be less than 1000 MW due to a lack of suitable space (Dena, 2005). The cost of producing wind power capacity has fallen drastically; the price of 1 KW wind energy capacity fell by 29% between 1990 and 2004 (Iset, 2005).

The layout of this model is as follows: a social planner or monopolist learn with some learning parameter λ, and it can choose any particular date in the future, when they would like to adopt a new technology. This is characterized by an increase of the learning parameter to γ, where $\gamma \geq \lambda$. After the innovation date, newly gained experience reduces the present level of cost by more, than before the innovation date. Thus the process innovation described here, is a substitute to learning. A firm can adopt a new technology when its research department has been successful. The cost of research is given by a convex and decreasing function $a(t_1)$, where t_1 is the date of innovation. When firms prefer an earlier innovation date over that of a late one, then they employ more researchers; this is reflected by a higher innovation cost in the model. The setting is deterministic to avoid unnecessary complications, which would not add any further results. Players choose the date of innovation at the beginning of the planning horizon; production starts thereafter. In the first step, a pricing rule is derived for an exogenous innovation date, which is endogenized thereafter.

The findings of this paper are: the social planner/ monopolist charges two different prices, for the time phases before the innovation and after the innovation. Both prices are constant for a constant price elasticity of demand. After the innovation has occurred, the decision maker's price rule, is such that the price (social planner) or the marginal revenue (monopolist) equal marginal cost at the last unit produced. This result is analogous to the findings of Spence (1981) who examined learning in the absence of innovations. However before the innovation occurs, the social planner's (monopolist's) price rule is such that the price (marginal cost) equals marginal cost at t_1 plus a negative constant. At the time of innovation, the costate variables of the two phases equal the ratio of the learning parameters λ/γ. Thus there is a downward jump in prices at the innovation date. In a second step, a subsidy on innovation cost and a subsidy on sales are introduced. The central results of these market interventions are: innovation subsidies and distribution subsidies reduce the prices of both phases if all variables (the timing of innovation t_1, the cumulative production quantities at the innovation date; $y(t_1)$ and at the end of the planning horizon $y(T)$) are endogenous. Both subsidy types induce innovation to proceed earlier. Consequently the total quantity produced during the entire planning horizon increases. The production plan in the presence of subsidies lies entirely above the production plan without subsidies. The result being, if early distribution yield positive externalities, then subsidies on sales and on innovation contain an additional positive effect.[3]

It is also shown that a subsidy on innovation cost (sales), which is financed through a tax on sales (innovation cost) changes the proportion of consumer and producer rents. Customers generally benefit more from sales subsidies, producers from innovation subsidies. Another central result emerges, when the total production $y(T)$ is restricted. In this case the two kinds of subsidies that are analyzed have different effects.

The next section introduces the model and solves for an optimal pricing rule, which is analyzed in detail. Section 3 endogenizes the timing of innovation. Section 4 continues with a welfare analysis. Section 5 concludes.

2 The Model with an exogenous innovation date

This model is solved for different market structures, at first the social planner's problem is solved, which can be easily extended to account for a market with perfect competition that yields quite similar results. The learning by doing case without innovation has been examined similarly by Brueckner et al. (1983). Later a monopolist takes the place of the social planner. This scenario is more relevant to reality, because in an environment of innovations, patents guarantee that their holders are able to execute market power. It has been rarely observed that a state runs a public firm in a learning industry, nevertheless a social planner's actions are examined as though they are almost identical to those of a monopolist.

2.1 The Social Planner's Problem

Assume there is a publicly owned firm, which faces the demand function: $x(p(t), t)$ for a non-storable output $x(t)$ that is sold at a price $p(t)$. Time is denoted by $t \in R^0_+$. The beginning of the first phase, when the planning horizon begins is t_0. The time when the innovation takes place is t_1. It is the end of the first phase and the beginning of the second phase. The planning horizon ends at $t = T$. The firm chooses an optimal time path for its control variables during the first phase, $p_0(t)$ and the second phase, $p_1(t)$; where $p(t) = \{p_0(t), p_1(t)\}$ The instantaneous production flows of the first and second phase are $x_0(p_0(t), t)$ and $x_1(p_1(t), t)$ respectively. They are the derivatives of the state variables $y_0(t)$ and $y_1(t)$, which are the cumulative production quantities for a period t before and after the innovation. Over the intervals $[t_0, t_1]$ and $[t_0, T]$, the social planner receives a stream of consumption benefits discounted back to $t = t_0$,

[3]For technologies in the renewable energy sector holds that early distributions increase their positive externality on the atmosphere more. The total carbon dioxide emissions are reduced more, because renewable energy sources can substitute conventional $C0_2$ emitting ones earlier.

$$\int_{t_0}^{t_1} B_0(p_0(t),t)e^{-r(t-t_0)}dt \text{ and } \int_{t_1}^{T} B_1(p_1(t),t)e^{-r(t-t_0)}dt \tag{1}$$

where $B_0(p_0(t),t)$ and $B_1(p_1(t),t)$ denote the per-period social surplus during the first and second phase. They are each equivalent to the area below the inverse demand function at some t, before and after t_1, respectively. The social planner faces a marginal cost that consists of two parts; a fixed part denoted by the parameter m, and a variable part that is equal to c at the beginning of the first phase, when experience $y_0(t)$ equals zero. This variable part decreases with a learning parameter λ before an innovation occurs. Intuitively there is continuous discounting involved, which is expressed by the exponential term.

$$MC_0(y_0(t)) = m + ce^{-\lambda y_0(t)} \qquad\qquad for\ t_0\ \leq t \leq t_1 \tag{2}$$

The social planner can adopt a new technology, when the research department has been successful. The faster an innovation occurs, the more costly it is. For now, the innovation cost function depends solely on the innovation date t_1. When no innovation occurs and the firm produces with the same technology during the entire planning horizon, then the innovation cost is zero; $a(t_1) > 0, \nabla t \setminus t = T$ where $a(T) = 0$. $a' < 0$, $a'' > 0$. A new technology is adopted right after the innovation. Otherwise, if a later date of innovation is preferred, the planner could reduce its cost by devoting fewer resources to its research department. A different cost function is introduced in section 2.4. The innovation cost is assumed to be paid in advance at t_0. After t_1 the firm faces more intensive learning; it learns with a learning parameter $\gamma \geq \lambda$. Experience completely transfers to the new technology . Switching costs are ignored, because they do not yield results, which extend the knowledge of the existing literature (see Parente, 1993). The second phase's marginal costs are

$$MC_1(y_1(t)) = m + ce^{-\lambda y_0(t_1)-\gamma[y(t)-y_0(t_1)]} \qquad\qquad for\ t_1\ < t \leq T \tag{3}$$

Thus the social planner's objective is,

$$\underset{p_0(t),\ p_1(t),\ t_1}{Max} SP \equiv \int_{t_0}^{t_1} e^{-r(t-t_0)} \left\{ B_0(p_0(t),t) - (m + ce^{-\lambda y_0(t)}) x_0[p_0(t),t] \right\} dt \tag{4}$$

$$-a(t_1)e^{-rt_0} + \int_{t_1}^{T} e^{-r(t-t_0)} \left\{ B_1(p_1(t),t) - (m + ce^{-\lambda y_0(t_1)-\gamma[y(t)-y_0(t_1)]}) x_1[p_1(t),t] \right\} dt,$$

where $\frac{\partial B_i}{\partial p_i} = p_i(t)\frac{\partial x_i[p_i(t),t]}{\partial p_i(t)}$ for $i \in (0,1)$. The constraints of the problem are given by

$$\dot{y_0}(t) = x_0[p_0(t),t] \quad t \in [t_0, t_1] \tag{5}$$

$$\dot{y_1}(t) = x_1[p_1(t),t] \quad t \in [t_1, T] \tag{6}$$

$$y_0(t_0) = 0 \tag{7}$$

$$y_0(t_1) = y_1(t_1) = y(t_1) \text{ is free} \tag{8}$$

$$y_1(T) \text{ is free} \tag{9}$$

$\dot{y_0}(t)$ and $\dot{y_1}(t)$ are time derivatives of cumulative production quantities or experience stocks. To keep this analysis simple, a real interest rate of zero is assumed. In the appendix it is shown, how the equilibrium changes when $r \neq 0$. Flows are functions of the price and time, where the price itself is a function of time. The cumulative quantity cannot change over night, when the innovation takes place and the new production process is adopted (8). Condition (9) is used as a transversality condition for the second phase. Necessary conditions of this problem are derived in two steps. Firstly this study examines some innovation date $t_1 \in [t_0, T]$ and solves for the price paths $p_0(t)$ and $p_1(t)$ with t_1 being fixed. In the next step the innovation date is endogenized.

Proposition 1 *A social planner chooses a constant price for each period of phase one and two respectively. The two prices are different across phase one* $[t_0, t_1]$ *and phase two* $[t_1, T]$.

Proof By a theorem of Hestens, take SP (4) with a fix t_1 and define $\eta_0(t)$ on the interval $[t_0, t_1]$ and $\eta_1(t)$ on the interval $[t_1, T]$ as the costate variables of the cumulative quantities $y_0(t)$ and $y_1(t)$ respectively.[4] The innovation cost function is $a(t_1)$. It can be ignored during the time the pricing rule is analyzed, because t_1 is fixed. Thus $a(t_1)$ is constant and drops out of the first order condition that describes the optimal pricing rule. The Hamiltonian is

$$H\ [p_0(t),\ p_1(t),\ \eta_0(t),\ \eta_1(t)] = B_0(t) - C_0(t) + B_1(p_1(t), t) - C_1(t) \tag{10}$$

$$-a(t_1) +\ \eta_0(t)x[p_0(t), t] + \eta_1(t)x[p_1(t), t]$$

where $C_i(t) = x_i(t)MC_i(y_i(t))$ for $i \in (0, 1)$ is the per-period cost. $p_0^*(t)$ and $p_1^*(t)$ maximize (10) such that

$$H\ [p_0^*(t),\ p_1^*(t),\ \eta_0(t),\ \eta_1(t)] \geq H\ [p_0(t),\ p_1(t),\ \eta_0(t),\ \eta_1(t)] \tag{11}$$
$$\text{for all } \ p_0(t) \geq 0,\ p_1(t) \geq 0.$$

The pricing rule for the first phase ($t \leq t_1$)
As $p_0^*(t)$ maximizes H for $(t \leq t_1)$, the necessary condition is

$$\frac{\partial H}{\partial p_0(t)} = p_0(t)\frac{\partial x_0(p_0(t), t)}{\partial p_0(t)} - \frac{\partial x_0(p_0(t), t)}{\partial p_0(t)}(m+ce^{-\lambda y_0(t)}) + \eta_0(t)\frac{\partial x_0(p_0(t), t)}{\partial p_0(t)} \doteq 0$$

[4] see Takayama p.658

$$\Leftrightarrow p_0(t) = m + ce^{-\lambda y_0(t)} - \eta_0(t) \tag{12}$$

The social planner sets a price that equals the marginal cost minus the shadow price of cumulative quantity at some t. The second necessary condition is

$$\dot{\eta_0}(t) = -\frac{\partial H}{\partial y_0(t)}$$

$$\Leftrightarrow \eta_0(t) = ce^{-\lambda y_0(t)} + const_1 \tag{13}$$

The third necessary condition is (5).

Lemma 2 *The shadow price at the end point of the first phase equals* $\eta_0(t_1) = \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)} - \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)}$.

Proof. See appendix. ∎

The second necessary condition (13) can be solved for $const_1$ with the transversality condition $\eta_0(t_1) = -\frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} + \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)}$. Evaluating $\eta_0(t)$ at t_1

$$const_1 = -\frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} + \frac{\lambda-\gamma}{\gamma}ce^{-\lambda y(t_1)}$$

$$\Longrightarrow \eta_0(t) = ce^{-\lambda y_0(t)} - \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} + \frac{\lambda-\gamma}{\gamma}ce^{-\lambda y(t_1)} \tag{14}$$

(13) and (14) solve for the price of the first phase

$$p_0 = m + \frac{\gamma-\lambda}{\gamma}ce^{-\lambda y(t_1)} + \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)}; \qquad t_0 \le t \le t_1 \tag{15}$$

During the first phase, p_0 is independent of time, which completes the first part of the proof of 1

The pricing rule for the second phase ($t > t_1$)

The necessary first order condition with respect to $p_1(t)$ can be solved for a function of the the second phase's costate

$$\frac{\partial H}{\partial p_1(t)} = p_1(t)\frac{\partial x_1(p_1(t),t)}{\partial p_1(t)} - \frac{\partial x_1(p_1(t),t)}{\partial p_1(t)}(m+ce^{-\lambda y(t_1)-\gamma[y(t)-y(t_1)]}) + \eta_1(t)\frac{\partial x_1(p_1(t),t)}{\partial p_1(t)} \doteq 0$$

$$\Leftrightarrow p_1(t) = m + ce^{-\lambda y(t_1)-\gamma[y(t)-y_1(t_1)]} - \eta_1(t) \tag{16}$$

The social planner's price is equal to the marginal cost minus the shadow price of cumulative quantity. The second condition that needs to be fulfilled is

$$\dot{\eta_1}(t) = -\frac{\partial H}{\partial y_1(t)}$$

$$\Leftrightarrow \eta_1(t) = ce^{-\lambda y(t_1)-\gamma[y(t)-y(t_1)]} + const_2 \tag{17}$$

The third condition is given by (6). $\eta(T) = 0$, because the value of experience at the end of the second phase is zero. The cumulative quantity at the end of the second phase is not restricted, hence (9) can be used to set up the following transversality condition, which solves for $const_2$.

$$const_2 = -ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)}$$

$$\Longrightarrow \eta_1(t) = ce^{-\lambda y(t_1)-\gamma[y(t)-y(t_1)]} - ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} \tag{18}$$

(17) and (18) are used to express the second phase's price

$$p_1 = m + ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)}; \qquad t_1 \le t \le T \tag{19}$$

For any t where $t_1 \le t \le T$, the price of phase 2 is constant. This completes the second part of proposition 1's proof.
It is assumed that the demand function does not change over time, thus the planner produces the same quantity in each period within the first phase and the same quantity within the second phase. The intuition behind this result is: although costs decrease through time, which would yield lower prices in a static model, the decrease of costs is completely offset by the decrease of the experience value in this dynamic framework. When either $\gamma = \lambda$ or $t_1 = T$, then (15) and (19) are equal: $p_0(t) = p_1(t) = m + ce^{-\lambda y(T)}$. For $\gamma > \lambda$ or $t_1 < T$, the price of the first phase exceeds the price of the second, which is as follows

$$p_0 > p_1$$

$$\Leftrightarrow \frac{\lambda}{\gamma}e^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - \frac{\lambda-\gamma}{\gamma}e^{-\lambda y(t_1)} > e^{(\gamma-\lambda)y(t_1)-\gamma y(T)}$$

$$\Leftrightarrow \frac{\lambda-\gamma}{\gamma}e^{(\gamma-\lambda)y(t_1)-\gamma y(T)} > \frac{\lambda-\gamma}{\gamma}e^{-\lambda y(t_1)}$$

$$\gamma[y(t_1) - y(T)] < 0$$

As the last expression holds, it follows that the claim $p_0 > p_1$ is correct. A social planner encounters a loss, because during the first phase, the price is below the marginal cost at t_1, during the second phase the price just covers its cost at $t = T$ and is below that level for all preceding periods. Therefore one would need to introduce a tax on a different market to compensate for the loss. The monopolist's problem is solved, before the results, which are quite similar are interpreted further.

2.2 The Monopolist's Problem

The monopolist's instantaneous profit functions for the two phases are,

$$\pi^0(t) \equiv \left[p_0(t) - (m + ce^{-\lambda y_0(t)})\right] x_0(p_0(t), t): \qquad t \in [t_0, t_1] \tag{20}$$

$$\pi^1(t) \equiv \left[p_1(t) - (m + ce^{-\lambda y_0(t_1)-\gamma[y(t)-y_0(t_1)]})\right] x_1(p_1(t), t): \qquad t \in [t_1, T] \tag{21}$$

where the variables and parameters are defined and interpreted in the social planner's problem. The firm's objective is,

$$Max\ MP = \int_{t_0}^{t_1} \pi^0(t)e^{-r(t-t_0)}dt - a(t_1)e^{-rt_0} + \int_{t_1}^{T} \pi^1(t)e^{-r(t-t_0)}dt \tag{22}$$

subject to constraints (5) to (9). In the absence of discounting, the Hamiltonian equals

$$H\ \ [p_0(t),\ p_1(t),\ \eta_0(t),\ \eta_1(t)] = \pi^0(t)+\pi^1(t)-a(t_1)+\ \eta_0(t)x[p_0(t),t]+\eta_1(t)x[p_1(t),t] \tag{23}$$

One can use the same methods that were used to derive (15) and (19) to derive the pricing rules when a monopolist is the decision maker

$$p_0\left(1-\frac{1}{\varepsilon(t)}\right) = m + ce^{-\lambda y(t_1)} - \eta_0(t)$$

$$\text{where } \varepsilon(t) = -\frac{\partial x_0}{\partial p_0}\frac{p_0}{x_0}$$

$$\Rightarrow MR_0(t) = m+ce^{-\lambda y(t_1)}-\left[\frac{\lambda}{\gamma}ce^{-\lambda y(t_1)} - ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)}\right]; \qquad t_0 \le t \le t_1 \tag{24}$$

(24) is the pricing rule before the innovation date,

$$p_1\left(1-\frac{1}{\varepsilon(t)}\right) = m + ce^{-\lambda y(t_1)-\gamma[y(t)-y_1(t_1)]} - \eta_1(t)$$

$$\text{where } \varepsilon(t) = -\frac{\partial x_1}{\partial p_1}\frac{p_1}{x_1}$$

$$\Rightarrow MR_1(t) = m + ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} \tag{25}$$

and (25) the innovation date after t_1. Therefore the monopolist sets the price where the marginal revenue equals marginal cost minus the shadow price of cumulative quantity. The only difference to the social planner's problem is that the optimal rule contains the multiplier $(1-\frac{1}{\varepsilon(t)})$, and thus the marginal revenue and not the price, appears in the optimality condition. Prices in the monopoly model are constant for constant elasticities $\varepsilon(t) = \varepsilon$, hence the same holds for the per period production quantities. When either the equality $\gamma = \lambda$ or $t_1 = T$ hold, then (24) reduces to (25): $MR_0(t) = MR_1(t) = m + ce^{-\lambda y(T)}$. This is the classical optimal pricing behavior of a learning monopolist in the absence of innovations shown by Spence (1981): "At every time, output should be profit maximizing output, given that marginal cost is the unit cost that obtains at the end of the period".[5] The total cost that a firm faces is the area

[5] See page 52.

underneath the learning curve or marginal cost curve between t_0 and T. If the firm increases output by ϵ_0, in any interval within $[t_0, t_1]$ or by ϵ_1, in any interval within $[t_1, T]$, then the incremental cost is not the cost that arises during that time. It is rather the change of the total area below the learning curve.

Proposition 3 *In the presence of innovations, a monopolist charges a price such that marginal revenue equals incremental cost at any point in time.*

Proof. See appendix. ■

Consequently, within the interval $[t_1, T]$ the monopolist prices optimally, when the marginal revenue at each point in time equals the marginal cost of the last unit produced. $MR_1(t) = MC|_{t=T}$ by (25). In the interval $[t_0, t_1]$ the monopolist charges $MR_0(t) = MC|_{t=t_1} - \frac{\lambda}{\gamma}\left[ce^{-\lambda y(t_1)} - ce^{-\lambda y(t_1)-\gamma[y(T)-y(t_1)]}\right]$ by (24), where the sum in brackets is positive. Consequently, the monopolist charges a lower price such that marginal revenue at each point in time is below the marginal cost of the last unit produced at t_1, because production continues beyond t_1. The "price discount" equals $\frac{\lambda}{\gamma}\left[ce^{-\lambda y(t_1)-\gamma[y(T)-y(t_1)]} - ce^{-\lambda y(t_1)}\right]$.[6] It contains information about how much the experience level $y(t_1)$ is worth for the production after t_1. In the next section this term is analyzed further.

If discounting is included in the analysis, then prices increase compared to those in (24) and (25) for all t. When future profits are discounted, then learning is valued less, because the experience payoff decreases. Thus in the presence of a positive discount rate, the firm increase its price over the entire planning horizon. In a model without discounting, learning is appreciated most in the beginning of the planning horizon, because its return lasts for a long period of time. In the absence of innovations the price difference between a model with and without a discount rate, reaches its peak at t_0. In this model, where innovation increases the learning parameter, the price difference could even be larger at t_1 than at t_0, because the learning intensity jumps. At T, prices that include discounting are equal to those where discounting is absent, because the return to experience is non-existing.

Proposition 4 *When $r \neq 0$, then a monopolist sets its price during the first phase, such that the following condition holds,* $MR_0(t) = m + \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} + \frac{\gamma-\lambda}{\gamma}ce^{-\lambda y(t_1)} +$

$$r\int_t^{t_1}\left[m + ce^{-\lambda y_0(\tau)}\right]e^{-r(\tau-t)}d\tau + r\int_{t_1}^{T}[m + ce^{(\gamma-\lambda)y(t_1)-\gamma y(\tau)}]e^{-r(\tau-t)}d\tau.$$

During the second phase the optimality condition is $MR_1(t) = m + ce^{(\gamma-\lambda)y(t_1)-\gamma y_1(T)} +$

$$r\int_t^T\left[m + ce^{(\gamma-\lambda)y(t_1)-\gamma y_1(\tau)}\right]e^{-r(\tau-t)}d\tau$$

Proof. See appendix. ■

The additional terms on the right hand side are positive, which implies that the price increases. At $r = 0$, the conditions reduce to (24) and (25).

[6] To be precise, the price discount also contains the constant multiplier $\left(1 - \frac{1}{\varepsilon}\right)^{-1}$, which is ignored in the following partial analysis.

The market equilibrium of a social planner and a monopolist are quite similar. In the past, infant industries have been heavily subsidized by governments, but they were not run as public firms. Examples are the aerospace and defense industry during and after World War Two, computer industries in the 1980/90s and firms that have operated in the renewable energy sector during the last 10 years. In the presence of learning and innovations, where the later can be protected by property rights, there are either monopolies or oligopolies in the market. This holds true for all industries mentioned above: Airbus and Boeing (aerospace market), Microsoft and IBM (software and hardware) and the renewable energy sector, where for instance five producers have a market share of over 90% of worldwide wind turbine sales.[7] Based on these real world observations, for the rest of this article, it seems reasonable to assume that a monopolist is the decision maker. Furthermore it does not matter much, because the pricing rules differ by a multiplier that depends on the demand elasticity.

2.3 A partial comparative analysis

The price discount of the first phase

The first phase's price discount is $\frac{\lambda}{\gamma}\left[ce^{-\lambda y(t_1)} - ce^{-\lambda y(t_1)-\gamma[y(T)-y(t_1)]}\right]$. This section studies the discount's size based on the underlying parameters. It follows a comparative analysis; the discount is partially differentiated with respect to the parameters λ, γ and c. It is important to note that all parameters affect the three variables t_1, $y(t_1)$ and $y(T)$, which are fixed here. This analysis is meant to explain intuitively the results that are derived later, when t_1 and $y(t_1)$ are endogenous, but $y(T)$ is not.

$$\frac{\partial(\cdot)}{\partial\lambda}[ce^{-\lambda y(t_1)} - ce^{-\lambda y(t_1)-\gamma[y(T)-y(t_1)]}]\left(\frac{1}{\gamma} - \frac{\lambda}{\gamma}y(t_1)\right) \tag{26}$$

The first bracket of (26) is positive, the second is positive for $\lambda y(t_1) < 1$, which is satisfied in the numerical simulation later. The price discount of the first phase rises with the learning parameter of the first phase, λ. The monopolist reduces its first phase's price to exploit a larger learning intensity.

$$\frac{\partial(\cdot)}{\partial\gamma} = -\frac{\lambda}{\gamma^2}\left[ce^{-\lambda y(t_1)} - ce^{-\lambda y(t_1)-\gamma[y(T)-y(t_1)]}\right] + \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)-\gamma[y(T)-y(t_1)]}\left[y(T) - y(t_1)\right] \tag{27}$$

In (27), the first summand is negative, because its bracket term is positive. The second summand is positive. The first summand exceeds the second in absolute value conditional on $e^{\gamma[y(T)-y(t_1)]} - 1 > \gamma[y(T) - y(t_1)]$. This condition is met when the produced quantity after t_1 is large enough. A large learning parameter after the innovation, γ decreases the incentive to reduce the incentive to reduce cost before the innovation date.

[7] Press release of BTM Consult ApS (27.3.2008).

$$\frac{\partial(\cdot)}{\partial c} = \frac{\lambda}{\gamma}\left[e^{-\lambda y(t_1)} - e^{-\lambda y(t_1)-\gamma[y(T)-y(t_1)]}\right] \tag{28}$$

(28) shows, how the variable part of the marginal cost level affects the price discount of the first phase.[8] The derivative is positive, because the return to experience increases when the original cost level is high. The price discount on p_0 increases with λ and c, it decreases with γ.

The Costates

The costate variables are positive for all t, however they decrease. $\eta_0(t)$ declines at a rate of the marginal cost's derivative for the first phase, $\eta_1(t)$ at a rate of the marginal cost's derivative for the second phase. An interesting result is that the quotient of the two costates at the optimal innovation time t_1 is the quotient of the learning parameters:

$$\eta_0(t_1) = \frac{\lambda}{\gamma} ce^{-\lambda y(t_1)} - \frac{\lambda}{\gamma} ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} = \frac{\lambda}{\gamma}\eta_1(t_1) \Rightarrow \frac{\eta_0(t_1)}{\eta_1(t_1)} = \frac{\lambda}{\gamma} \tag{29}$$

Figure 1 shows the course of two costates, given a cumulative production quantity at the innovation date of 50 and 100. The cumulative quantity at the end of the planning horizon is 150, thus experience becomes worthless and both costate functions converge to the horizontal axis. At $y(t_1)$ the costates jump vertically upward such that the costate at $\lim\limits_{\epsilon\to 0}\,[y(t_1)+\epsilon]$ is $\frac{\gamma}{\lambda}$ times larger compared to its value at $\lim\limits_{\epsilon\to 0}\,[y(t_1)-\epsilon]$. Prices have been derived as functions of costates during phase one, $(1-\frac{1}{\varepsilon})^{-1}p_0(t) = m+ce^{-\lambda y(t)} - \eta_0(t)$ [see (24)] and phase two, $(1-\frac{1}{\varepsilon})^{-1}p_1(t) = m+ce^{-\lambda y(t_1)-\gamma[y(t)-y_1(t_1)]} - \eta_1(t)$ [see (25)]. At t_1, the prices $p_0(t)$ and $p_1(t)$ reduce to $\left(m+ce^{-\lambda y(t_1)}\right)(1-\frac{1}{\varepsilon})$, subtracted by $\eta_i(t)(1-\frac{1}{\varepsilon})$ for $i \in (0,1)$. Figure 1 clearly shows an upward jump of costates, which means that prices drop discontinuously by the amount that the costates jump with their constant multiplier.

In the past there have been government interventions that aimed to sell a fix number of products, which are characterized through positive externalities e.g. solar panels.[9] Figure 1 shows a decrease of $y(t_1)$ from 100 to 50, keeping $y(T)$ and all parameters constant. The costate of the function, where the innovation occurs earlier is lower during its first phase compared to the other costate. It exceeds the other costate thereafter before $y(t_1) = 75$ is reached. Afterwards it is lower again. Therefore the price during the first phase decreases with $y(t_1)$. The next proposition shows that the same holds true for t_1.

Proposition 5 *An earlier innovation date t_1 increases the price p_0 and decreases the price p_1, iff $y(T)$ is fixed.*

[8] Recall that $MC|_{t=0} = m + c$

[9] The "100,000 roof-program" was part of the Renewable Energy Law in Germany. It intended to install 100,000 solar panels (which would be equivalent to $y(T)$) in a given time (T).

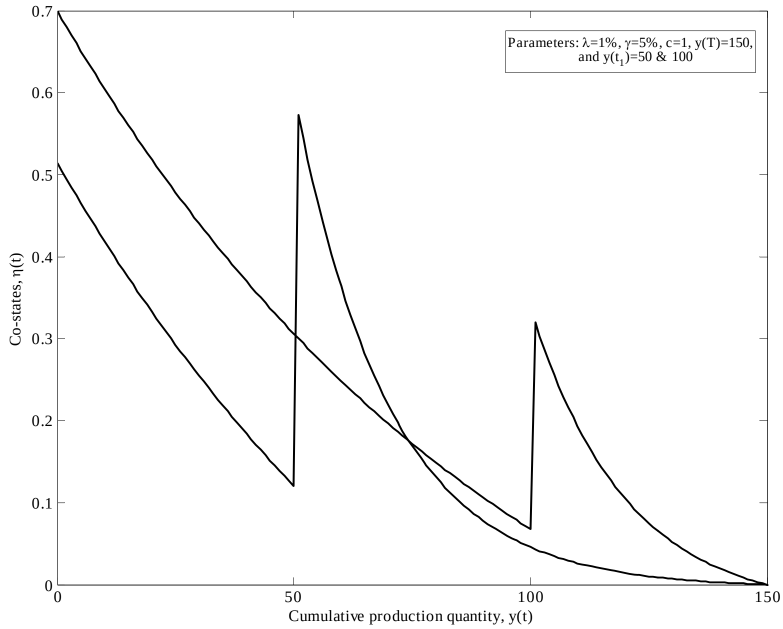

Figure 1: Costates jump at the date of innovation

Proof. The elasticity of demand is assumed to be constant and the monopolist chooses the optimal innovation date t_1 before production starts. Given p_0 (24) and p_1 (25), prices are differentiated with respect to t_1.

$$\frac{\partial p_0}{\partial t_1} = \underbrace{\frac{(\gamma-\lambda)\lambda}{\gamma}}_{>0} c \underbrace{\left[e^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - e^{-\lambda y(t_1)}\right]}_{<0} \underbrace{\frac{\varepsilon}{\varepsilon-1}}_{>0} \underbrace{\frac{y(t_1)}{\partial t_1}}_{>0} < 0 \qquad (30)$$

In (30) all terms except of one, are positive for $\gamma > \lambda$, $c > 0$, $y(t_1) < y(T)$. Later it is shown numerically that a delay of the innovation date increases the cumulative quantity up to the innovation date: $\frac{y(t_1)}{\partial t_1} > 0$.

$$\frac{\partial p_1}{\partial t_1} = \frac{\partial p_1}{\partial y(t_1)} \frac{y(t_1)}{\partial t_1} = c \underbrace{\left[(\gamma-\lambda)e^{(\gamma-\lambda)y(t_1)-\gamma y(T)}\right]}_{>0} \underbrace{\frac{\varepsilon}{\varepsilon-1}}_{>0} \underbrace{\frac{y(t_1)}{\partial t_1}}_{>0} > 0 \qquad (31)$$

The price after t_1 increases, when the innovation occurs earlier. ■

This result seems to be puzzling at first glance. In the presence of learning without process innovations, the learning effect is smaller than the level effect of costs. Fudenberg and Tirole (1983) show that "output increases over time,

... [but] produce a lot now to lower costs, then ease off as an optimal control strategy" [was ruled out by their results].[10] This model also contains the same effects as in the Fundenberg and Tirole model: A firm chooses a production plan that maximizes today's profits taking account of all future cost reductions, where the later is determined by the learning effect. In the presence of innovations, the learning effect is stronger, when the date of innovation occurs later. Therefore a firm reduces p_0 when t_1 increases. At the same time it increases p_1, because future time $(T - t_1)$ decreases along with the benefit of future cost reduction. The "today's-profit maximizing-effect" is stronger than the learning effect and a decrease of t_1 comes with an increase of p_0. The intuitions provided by this partial analysis are helpful for section 3, where t_1^* is endogenous. The numerical part of section 3 contains two parts; in the first part all variables are endogenous, and it is shown that (30) and (31) do not hold. In the second part, where t_1^* and $y(t_1^*)$ are endogenous, but $y(T)$ is not, the results provided here, appear again.

2.4 A different innovation cost function

In this section the thus far neglected innovation cost function with a different structure is introduced. It was able to be neglected, because it depended solely on the innovation date, which was exogenous. What happens, when one assumes that the innovation cost function also depends on the experience accumulated before t_1 and $a = a[y(t_1), t_1]$, with $\frac{\partial a}{\partial y(t_1)} < 0$? This assumption is reasonable, when the research department works closely together with the production floor. The altered transversality condition of the first phase is $\eta_0(t_1) = -\frac{\lambda}{\gamma} c e^{(\gamma-\lambda)y(t_1)-\gamma y(T)} + \frac{\lambda}{\gamma} c e^{-\lambda y(t_1)} - \frac{\partial a[y(t_1),t_1]}{\partial y(t_1)}$. As innovation cost decreases with $y(t_1)$, so does the marginal revenue

$$MR_0(t) = m + \frac{\gamma - \lambda}{\gamma} c e^{-\lambda y(t_1)} + \frac{\lambda}{\gamma} c e^{(\gamma-\lambda)y(t_1)-\gamma y(T)} + \frac{\partial a[y(t_1), t_1]}{\partial y(t_1)}; \qquad t_0 \leq t \leq t_1 \tag{32}$$

Accordingly the price during the first phase decreases. The intuition behind the changes are obvious; experience does not only reduce future production cost, but also the cost of research.

3 The model with an endogenous innovation date

3.1 Analytical part

The optimal timing of innovation has been exogenous thus far, in this section it will be endogenized. First the problem is solved for an innovation cost function of the form $a = a(t_1)$. Later it is shown, what changes when $a = a[y(t_1), t_1]$. With t_1 fixed, consider the following problem with the two segments,

$$S.0. \max_{p_0(t)} \int_{t_0}^{t_1} \{x_0(p_0, t)(p_0 - m - c e^{-\lambda y_0(t)})\} dt \tag{33}$$

[10] See proposition 2 on p. 525.

$$s.t.\dot{y_0}(t) = x_0(p_0, t),\ t \in [t_0, t_1] \quad t_0, t_1 \text{ fixed}$$

and

$$S.I.\max_{p_1(t)} \int_{t_1}^{T} \{x_1(p_1, t)(p_1 - m - ce^{(\gamma-\lambda)y(t_1) - \gamma y_1(t)})\}dt \tag{34}$$

$$s.t.\dot{y_1}(t) = x_1[p_1, t],\ t \in [t_1, T] \quad t_1, T \text{ fixed}$$

When p_0^* and p_1^* are solutions to the problem MP (22), given conditions (5) to (9), then with the Hamiltonians being defined as usual,

$$H_0(p_0^*,\ \eta_0(t)) \geq H_0(p_0,\ \eta_0(t)) \qquad \nabla p_0(t) \geq 0 \tag{35}$$

$$H_1(p_1^*,\ \eta_1(t)) \geq H_1(p_1, \eta_1(t)) \qquad \nabla p_1(t) \geq 0 \tag{36}$$

where

$$H_0(p_0,\ \eta_0(t)) = \pi^0(t) + \eta_0(t)x(p_0, t)$$

$$H_1(p_1,\ \eta_1(t)) = \pi^1(t) + \eta_1(t)x(p_1, t)$$

Adding (35) and (36), one can see that p_0^* and p_1^* also satisfy Hamiltonian condition (11). Therefore the control variables that solve (22) also solve problems (35) and (36) respectively. Denote the maximized values of the objectives $S.0.$ (33) by $V_0^*(t)$ and $S.I.$ (34) by $V_1^*(t)$. A standard result of optimal control theory is

$$\frac{\partial V_0^*(t)}{\partial t_1} = H_0(t_1) \text{ and } \frac{\partial V_1^*(t)}{\partial t_1^*} = -H_1(t_1^*) \tag{37}$$

Consider the optimal value of t_1^*, denoted by $t_1^* \in (t_0, T)$. If t_1^* is optimal, it must solve

$$\max_{t_1} \left\{ \int_{t_0}^{t_1} \pi^0(p_0, t)dt + \int_{t_1}^{T} \pi^1(p_1, t)dt - a(t_1) \right\} = V_0^*(t) + V_1^*(t) - a(t_1) \tag{38}$$

$$\frac{\partial}{\partial t_1}[V_0^*(t) + V_1^*(t) - a(t_1)] = H_0(t_1^*) - H_1(t_1^*) - \frac{\partial a(t_1^*)}{\partial t_1^*} \doteq 0$$

$$\Leftrightarrow a'(t_1^*) = \pi^0(t_1^*) + \ \eta_0(t_1^*)x[p_0, t_1^*] - \pi^1(t_1^*) - \eta_1(t_1^*)x[p_1, t_1^*] \tag{39}$$

(39) reveals that the optimal innovation date t_1^*, the difference of the first and second phase's returns equal the derivative of the innovation function with

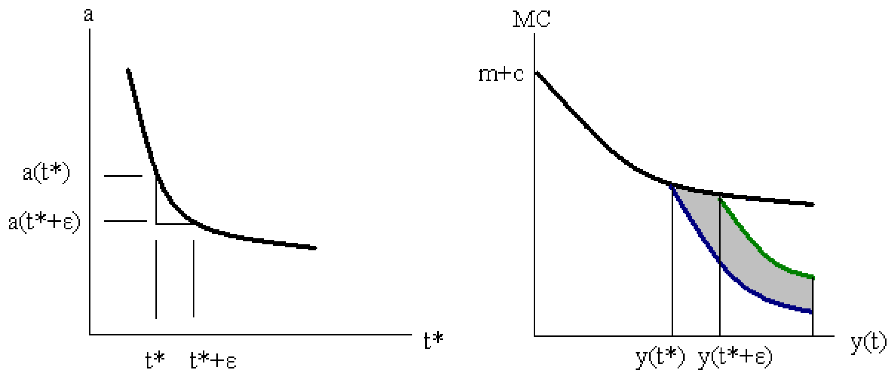

Figure 2: Benefits and costs of delaying innovation

respect to t_1^*, where the word 'return' circumscribes the instantaneous profit flow $\pi(t_1^*)$ and the return to experience $\eta(t_1^*)x[p, t_1^*]$. A simple innovation cost function $a(\cdot)$ that shall mimic reality has the following characteristics: $a(t_1) > 0$ $\nabla t \setminus t = T$ where $a(T) = 0$, $a' < 0$, $a'' > 0$. Thus innovation is costly if it is implemented within the planning horizon. The cost is proportionally larger, the sooner innovation takes place. The right side of (39) is moderately negative over the entire planning horizon, which can be shown numerically. Therefore the equality of both sides is guaranteed for an appropriate set of parameters. Figure 2 demonstrates that postponing innovation comes along with lower innovation cost (left diagram), but an increase of the production cost (on the right). The optimality condition for t_1^* (39) conveys the same result; in the optimum the cost savings of delaying innovation per period (left side) equals the differences of return of phase 1 and 2 (right side).

In section 2.4, the innovation cost function was changed to $a[y(t_1), t_1]$. If one accounts for these changes here, (39) adjusts to

$$\frac{\partial a[y(t_1^*), t_1^*]}{\partial t_1^*} + \frac{\partial a[y(t_1^*), t_1^*]}{\partial y(t_1^*)} \frac{\partial y(t_1^*)}{\partial t_1^*} = \pi^0(t_1^*) + \eta_0(t_1^*)x(p_0) - \pi^1(t_1^*) - \eta_1(t_1^*)x(p_1, t_1^*) \tag{40}$$

If $\frac{\partial a[y(t_1), t_1]}{\partial t_1}$ has not changed and the assumptions $\frac{\partial a[y(t_1), t_1]}{\partial y(t_1)} < 0$ and $\frac{\partial y(t_1)}{\partial t_1} > 0$ hold, then the left side of (40) decreases. The optimal date of innovation would increase and the monopolist's research budget would decrease. When an increase in $y(t_1)$ decreases the innovation cost, then the value of waiting with an innovation adoption increases.

3.2 Numerical part

After introducing a specific demand function, a numerical simulation method is used to find specific values of t_1^*, $y(t_1^*)$ and $y(T)$. Tax and subsidy parameters are added simultaneously. They are also helpful for the next section, when a welfare analysis is carried out.

Demand: The per period inverse demand function is,

$$p(x_i) = \frac{{x_i}^{-\alpha}}{1-\alpha} \tag{41}$$

for $i \in (0,1)$. For $p(x_i)$ as defined in (41), the price elasticity of demand is $\varepsilon = -\frac{\partial x}{\partial p}\frac{p}{x} = \frac{1}{\alpha}$.[11] It follows from (24) and (25) that $MR_i = k_i \Leftrightarrow p_i(1-\alpha) = k_i \Leftrightarrow x_i^{-\alpha} = k_i$, where

$$k_0 = \left(m + \frac{\lambda}{\gamma} c e^{(\gamma-\lambda)y(t_1^*)-\gamma y(T)} + \frac{\gamma-\lambda}{\gamma} c e^{-\lambda y(t_1^*)}\right)$$

$$k_1 = \left(m + c e^{(\gamma-\lambda)y(t_1^*)-\gamma y(T)}\right)$$

Sales subsidy/ tax: An ad valorem subsidy ($\tau > 1$) or tax parameter ($\tau < 1$) is added, a subsidy shifts the demand function out, a tax shifts it in. A demand function of the the type in (41), which includes the parameter τ is

$$x_i = \left(\frac{\tau}{p_i(1-\alpha)}\right)^{\beta} \tag{42}$$

It is helpful to transform (39) such that the t_2-optimality condition becomes

$$x_0 \left[p_0 - MC(t_1^*) + \eta_0(t_1^*)\right] - x_1 \left[p_1 - MC(t_1^*) + \eta_1(t_1^*)\right] = a'(t_1^*) \tag{43}$$

(29), (42) and (2) or (3) are substituted in (43)

$$\frac{\alpha\tau^{\beta}}{1-\alpha}\left(k_0^{1-\beta} - k_1^{1-\beta}\right) = a'(t_1^*) \tag{44}$$

Innovation cost: An innovation cost function that fulfills the requirements; $a(t_1^*) > 0, \nabla t \setminus t = T$ where $a(T) = 0$. $a' < 0$, $a'' > 0$ is

$$a(t_1^*) = ie^{-\mu t_1^*} - ie^{-\mu T} \tag{45}$$

The parameter i affects the level of the innovation cost, the parameter μ the slope with respect to t_1^*.

Innovation subsidy/ tax: The innovation subsidy is constructed in a way such that the government either pays a part of the innovation cost, $s > 0$ or charges a tax, $s < 0$. Thus the gross innovation cost is

$$a(t_1^*) = \rho[ie^{-\mu t_1^*} - ie^{-\mu T}] \tag{46}$$

[11] Define $\beta = \frac{1}{\alpha}$. One can easily account for a time-varying demand function by multiplying $p(x_i)$ with $b(t)$ where $b(t)^{\beta} = be^{\delta t}$. But this does not add much to this analysis.

where $\rho = 1 - s$. $a'(t_1^*)$ decreases with ρ, hence it becomes less beneficial for the firm to procrastinate the innovation timing and the innovation date occurs earlier when innovation is subsidized. Equation (44) can be rewritten as

$$\frac{\alpha \tau^\beta}{1-\alpha}\left(k_0^{1-\beta} - k_1^{1-\beta}\right) = -\rho \mu i e^{-\mu t_1^*} \tag{47}$$

(47) is the optimality conditions for t_1. In addition, one needs to express $y(t_1^*)$ and $y(T)$ to solve for these three variables simultaneously. The demand function (42) and the optimality condition $p_i = (1-\alpha)^{-1} k_i$ solve for the per period demand of phase 1, $x_0 = \left(\frac{\tau}{k_0}\right)^\beta$. As the per period production is constant, the cumulative production quantity up to the innovation date t_1^* is simply

$$y^M(t_1^*) = \left(\frac{\tau}{k_0}\right)^\beta (t_1^* - t_0) \tag{48}$$

Similarly, $y(T)$ equals the integral of the second phase's per period production flows between t_1^* and T, which is added to the cumulative quantity up to t_1^*

$$y^M(T) = y(t_1^*) + \left(\frac{\tau}{k_1}\right)^\beta (T - t_1^*) \tag{49}$$

Equations (47), (48) and (49) are three independent equations that contain as many unknowns t_1^*, $y(t_1^*)$ and $y(T)$. The parameters c, m, i, γ, λ, α, μ, δ and the planning horizon, t_0 and T are known to the firm. Thus the model can be solved numerically.[12]

Results

When $y(T)$ is endogenous then (30) does not hold. Figure 3 illustrates, how the variables of the model t_1^*, $y(t_1^*)$, $y(T)$ and the prices p_0, p_1 are affected by τ, ρ (both upper part), λ, and γ (both lower part). The vertical axis of the cumulative quantities is on the right side of each of the four panels, the vertical axis of all other variables on the left side. In the upper left panel sales taxes ($\tau < 1$) and subsidies ($\tau > 1$), in the upper right panel innovation subsidies ($\rho < 1$) and taxes ($\rho > 1$) vary. Taxes/ subsidies are absent when $\tau = \rho = 1$.[13] Taxes and subsidies, either on sales or innovation cost have the same effect. Taxes increase the prices of both phases and decrease the total cumulative production. The cumulative quantity at the innovation date increases, whenever innovation is postponed.

[12] So far a process innovation was analyzed, it is however fairly simple to account for a product innovation in this model. When a firm does not change its cost structure, but its product features, then it is possible to change the demand function from one with an exponent of α to one with an exponent different from α after t_1^*. This article considers the case, where consumers do not anticipate price changes, thus the demand is equal before and after t_1^*.

[13] The command 'fsolve' of the computer program MATLAB solves systems of nonlinear equations. It was used to derive t_1, $y(t_1)$ and $y(T)$ based on (47), (48) and (49). The parameters have been arbitrarily chosen. The results hold for other parameter sets, which yields a solution. The parameters used here are: $\alpha = 0.9$, $\gamma = 0.05$, $\lambda = 0.01$, $t_0 = 0$, $\mu = 0.1$, $m = 1$, $c = 5$, $T = 50$ and $i = 20$.

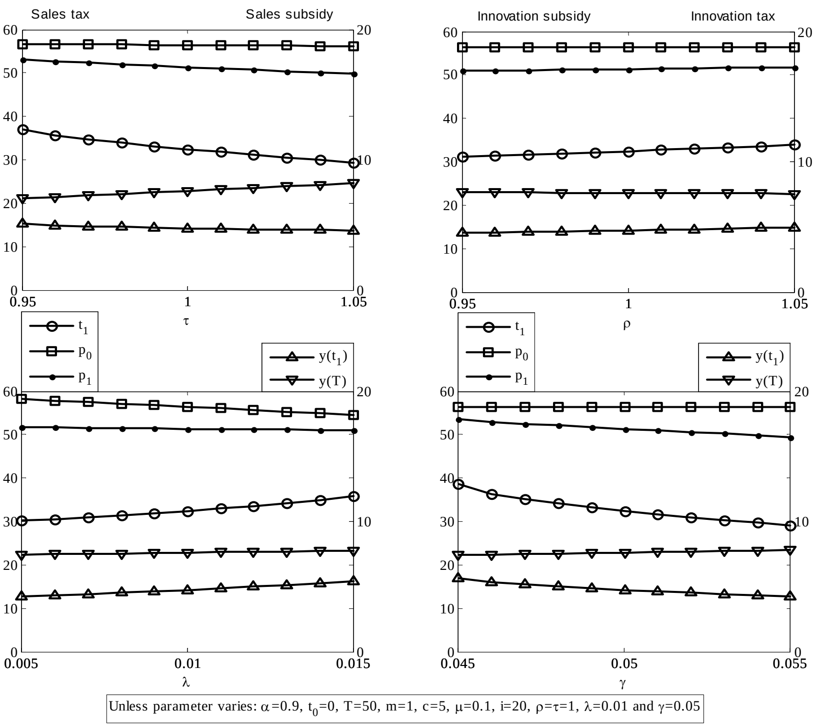

Figure 3: A comparative analysis, when all variables are endogenous

The lower panel demonstrates what happens when the learning parameter λ varies in a range of $[0.5\%, 1.5\%]$, the lower right panel shows γ varying in a range of $[4.5\%, 5.5\%]$. A rise in either learning parameter reduces both prices and increases total cumulative quantity. Innovation is delayed when λ increases, it occurs earlier, when γ increases. $y(t_1^*)$ moves into the same direction as t_1^*. Subsidies and more intense learning therefore are shown to reduce prices. A large λ decreases both prices but the larger price p_0 is charged by the monopolist longer, because a postponement of innovation means that the length of phase 1 increases. In the absence of discounting the average per period price is still smaller. But if consumer discounting is high, a large λ might not be beneficial for buyers.

So far positive externalities, which justify subsidies have been ignored. An early distribution of solar panels is preferred to a later date, to reduce the total carbon dioxide concentration in the atmosphere. Both types of subsidies yield a lower p_0 and a lower t_1^*, which together guarantee that with regards to time, more products are sold earlier, because $p_0 > p_1$. Thus both subsidies are beneficial, when early distributions play a role.

Results change drastically, when $\overline{y(T)}$ is given. The results of the partial analysis, (30) and (31) are supported by a numerical analysis, illustrated in figure 4. The structure and the parameters of this figure are the same as in figure 3. In figure 3 a total cumulative production quantity of around $y(T) = 8$ was derived. Figure 4 shows the results, when $y(T) = 8$ is assumed to be given exogenously. Thus the optimality condition (49) is excluded from the analysis. The results here have changed dramatically: innovation is postponed when sales are subsidized. It occurs earlier, when the innovation cost is subsidized. p_0 increases when innovation is subsidized, it decreases, when sales are subsidized. p_1 moves into the opposite direction. These results correspond with the analytical study above, where it was shown that $\frac{\partial p_0}{\partial t_1} < 0$ and $\frac{\partial p_0}{\partial t_1} > 0$ given $y(T)$ being exogenous. In the absence of subsidies the average production unit price equals $p^{AV} = y(t_1^*)p_0 + [y(T) - y(t_1^*)]\, p_1 = 53.47$. If a low sales subsidy ($\tau = 1.05$) is introduced, then the average price increases to $p^{AV,S} = 53.89$. An innovation subsidy ($\rho = 0.95$) decreases the average price $p^{AV,I} = 53.27$. If a regulator cares about consumer prices, but it wants to subsidize the industry as it yields positive externalities, then an innovation subsidy is preferred over a sales subsidy. Even though an innovation subsidy increases p_0, the average price decreases. In order to keep p_0 from rising, a price constraint could be implemented for the first time phase.

Proposition 6 *A price constraint during the first phase increases the consumer surplus and decreases* t_1^*.

Proof. See appendix. ■

The lower panels show the affect of the learning parameters. An increase of λ reduces p_0 strongly, while p_1 increases moderately. An increase of the second phase's learning parameter γ decreases p_1 and increases p_0. This observation is in line with the results from the partial analysis (26) and (27). A change of any learning parameter affects the timing of innovation, which influences the cumulative quantity at the innovation date positively if t_1^* increases and negatively if t_1^* decreases. A change of $y(t_1^*)$ affects the prices as shown by (30) and (31). The average price decreases, when either parameter increases. In the baseline case $\lambda = 1\%$, $\gamma = 5\%$ and the average price is $p^{AV} = 53.47$. When $\lambda = 1.5\%$, the average price increases to $p^{AV,\lambda} = 52.87$, when $\gamma = 5.5\%$ the average prices increases to $p^{AV,\gamma} = 52.67$.

When the timing of early distribution plays a role, one has to examine a change of t_1^* and p_0 to evaluate a policy or change of parameters in the same way as it has been done, when $y(T)$ was endogenous. An innovation subsidy yields a higher p_0, but innovation occurs earlier than in the presence of a sales subsidy, where p_0 is lower and t_1 is larger. With an innovation subsidy, fewer products are sold during the first phase. However, the second phase, during which the number of products sold per period is larger (as $p_1 < p_0$) begins earlier. Thus a regulator who is concerned about early distributions prefers a sales subsidy towards an innovation subsidy if she is very impatient.

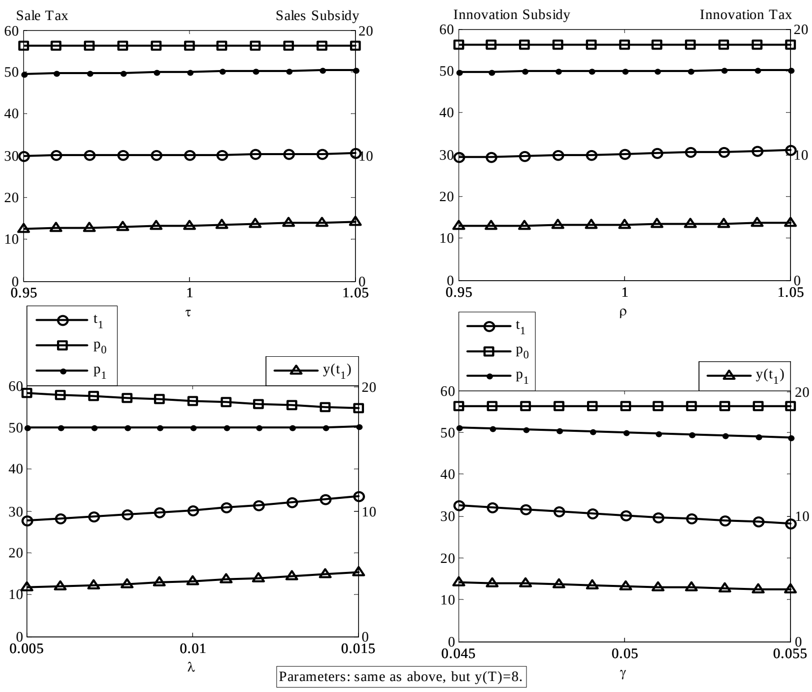

Figure 4: Comparative statics with $\overline{y(T)}$ being exogenously determined

4 More on welfare effects

This section contains a welfare analysis based on consumer and producer rents, which are first derived. It is illustrated that the two types of subsidies of either innovation or sales have different effects on consumer/ producer rents. This article does not contain a general welfare analysis, which needs to verify clearly the positive externalities that would induce a state to intervene. It would only be rational to do this, if one considers a specific industry with information pertaining to the model's underlying parameters, which is beyond the scope of this paper. This section shows that an innovation (sales) subsidy, which is fully financed through a sales (innovation) tax, changes the proportion of consumer and producer rents. This way consumer rents could be increased, which has a market power mitigating effect.[14] In this part all variables are endogenous. The same inverse demand function holds; $p = \frac{\tau}{1-\alpha}x^{-\alpha}$ and in the numerical analysis, the same set of parameters is used as before.

Consumer Rents: The per-period consumer rent during the first phase, CR_t^0 equals the area underneath the inverse demand function above the monopoly price.[15]

$$CR_t^0 = \int_{p_0}^{\infty} x_0[p_0(t)]dp = \left(\frac{\tau}{1-\alpha}\right)^{\beta} \int_{p_0}^{\infty} p^{-\beta}dp = \left(\frac{\tau}{1-\alpha}\right)^{\beta} \frac{-1}{1-\beta} p_0^{1-\beta}$$

$$= \frac{\alpha}{(1-\alpha)^2} k_0^{1-\beta} \tau^{\beta} \tag{50}$$

The price is substituted for $k_0/(1-\alpha)$ and $k_0 = m + \frac{\gamma-\lambda}{\gamma} ce^{-\lambda y(t_1^*)} + \frac{\lambda}{\gamma} ce^{(\gamma-\lambda)y(t_1^*)-\gamma y(T)}$.[16] The consumer rent over the entire first phase CR^0 is

$$CR^0 = \int_{t_0}^{t_1^*} CR_t^0 dt = \frac{\alpha \tau^{\beta} k_0^{1-\beta}}{(1-\alpha)^2}(t_1^* - t_0) \tag{51}$$

Applying the same steps again, the per period consumer rent after the innovation has taken place CR_t^1 and the consumer rent over all these periods, CR^1 are

$$CR_t^1 = \frac{\alpha \tau^{\beta} k_1^{1-\beta}}{(1-\alpha)^2} \tag{52}$$

$$CR^1 = \frac{\alpha \tau^{\beta} k_1^{1-\beta}}{(1-\alpha)^2}(T - t_1^*) \tag{53}$$

[14] A valuation of such a measure is not part of this analysis. It depends on the specific industry to judge, if such a procedure is justifiable.

[15] A necessary assumption is that the elasticity of demand exceeds one; $\beta = \frac{1}{\alpha} > 1$, to guarantee $\lim_{p\to\infty} p^{1-\beta} = 0$.

[16] An increase of α is equivalent to a decrease of the demand elasticity. Consider (50), an increase of α increases CR_t^0, which is quite intuitive. A monopolist picks a lower price when the elasticity is large.

where $k_1 = m + ce^{(\gamma-\lambda)y(t_1^*)-\gamma y(T)}$.
Producer Rents: The producer rent of the monopoly, which equals its profit over the entire planning horizon is

$$PR = \int_{t_0}^{t_1^*} x_0 \left(p_0 - ce^{-\lambda y(t)} - m\right) dt + \int_{t_1^*}^{T} x_1 \left(p_1 - ce^{(\gamma-\lambda)y(t_1^*)-\gamma y(t)} - m\right) dt$$

$$= \left(\frac{k_0}{1-\alpha} - m\right)\left(\frac{\tau}{k_0}\right)^{\beta}(t_1^* - t_0) + \left(\frac{k_1}{1-\alpha} - m\right)\left(\frac{\tau}{k_1}\right)^{\beta}(T - t_1^*)$$
$$+\frac{1}{\lambda}\left[ce^{-\lambda y(t_1^*)} - ce^{-\lambda y(t_0)}\right] + \frac{1}{\gamma}\left[ce^{(\gamma-\lambda)y(t_1^*)-\gamma y(T)} - ce^{-\lambda y(t_1^*)}\right] \quad (54)$$

Results

The results are summarized by figure 5. In the absence of government interventions consumer and producer rents are given by the horizontal line. After a sales subsidy is introduced to the market, consumer and producer rents jump upwards to the level, where the falling curves touch the vertical axis. Then an innovation tax comes into place. It is depicted on the horizontal axis of each panel. An increase of the innovation tax reduces consumer and producer welfare, which is expressed by the falling graph. Producer rents fall much faster than consumer rents. Hence a sales subsidy, which is financed through an innovation tax has market power mitigating effects. This illustration shall not propose such a market intervention, rather it shows that either subsidy type has different effects on producer and consumer rents. A similar figure could be shown for welfare changes through an innovation subsidy that is financed by a sales tax, where the upper curve is steeper for consumers.

5 Conclusion

The innovation of this model is that it is able to evaluate distribution and innovation subsidies, while innovation costs depend on time, and learning depends on cumulative production. This article examines the pricing of a monopolist and a public owned firm in an environment, where the unit cost of production decreases through learning. The learning intensity decreases with cumulative production. The firm can invest in an innovation process. At the time, when research is successful, learning for a given production quantity jumps. Thus a lower cost level can be achieved with less cumulative production added. So far models, which have included learning and innovation have assumed that learning occurs through time. Hence these models are not applicable to questions in the area of industrial organization, because learning has no effect on production or pricing.

In the absence of discounting it is shown that prices before and after the innovation date are constant. The price of the first phase exceeds the price of

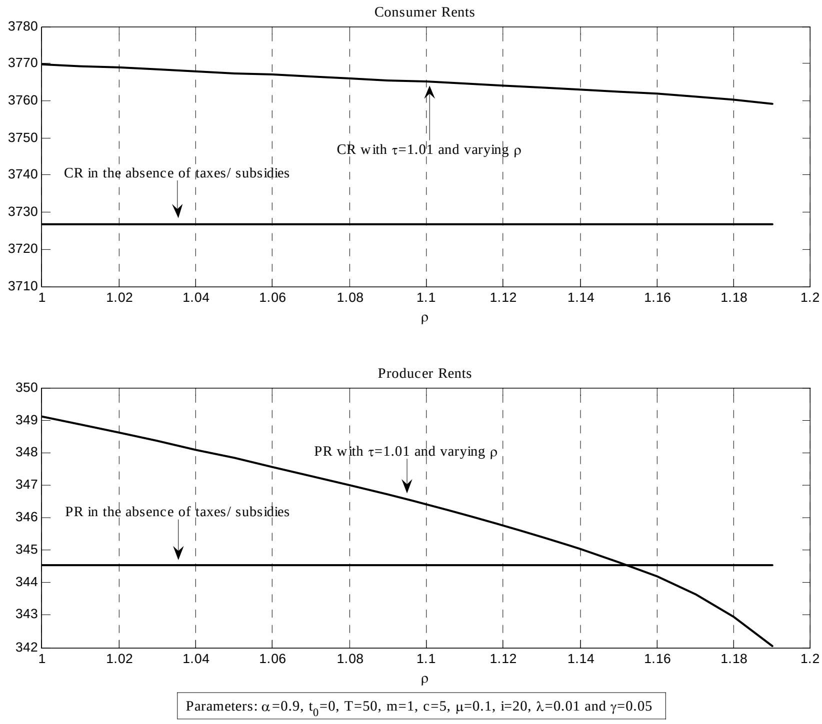

Figure 5: Welfare effects of government interventions

the second phase, hence there is a downward jump when the new technology is implemented. If discounting is included in the analysis, prices before and after the innovation date rise, because the return on learning falls. The problem of a monopolist and social planner are quite similar, however, this analysis concentrates on the monopolist, because learning industries, although often severely subsidized, are generally not publicly owned. Other central results of this paper are: innovation subsidies and sales subsidies reduce the prices of both phases if all variables (the timing of innovation t_1, the cumulative production quantities at the innovation date, $y(t_1)$ and at the end of the planning horizon, $y(T)$) are endogenous. Both types of subsidies induces the date of innovation to occur earlier. Therefore the total quantity produced during the entire planning horizon increases. The production plan in the presence of subsidies lies entirely above the production plan without subsidies. Thus if early distributions yield positive externalities, then subsidies on sales and on innovation contain an additional positive effect.[17] Another central result emerges, when the total production $y(T)$ is restricted.[18] In this case the two kinds of subsidies analyzed have different effects. Innovation subsidies decrease the innovation date, but the price during the first phase increases (thus sales decrease). It is shown that a price cap can reduce the first phase's price, furthermore it induces innovation to occur earlier. In order to evaluate an innovation subsidy one would need to consider the negative effect of the "early-distribution argument". Sales subsidies induce the innovation timing to occur later, which means that for a longer period of time, consumer pay the higher first phase's price p_0. The positive effect is that p_0 falls, which is why the "early-distribution argument" might be in favour of sales subsidies. The two subsidies considered have different effects on consumer and producer surplus. It is also shown that a subsidy on innovation (sales), which is financed through a tax on sales (innovation), changes the proportion of consumer and producer rents. Customers generally benefit more from sales subsidies, producers from innovation subsidies.

This paper does not conduct a general welfare analysis; for that one would need to measure the cause for government subsidies (positive externalities). This would only be feasible, if one would have specific learning parameters and the parameters of the innovation cost function, which describe a specific market. This goes beyond the scope of this article and is left for future work. It would also be interesting to show, how the market equilibrium changes, when the date of innovation is anticipated by customers. Most likely it would be optimal for some to wait and purchase the product after the innovation date. This would reduce the learning before t_1 unless innovation is delayed by the firm. Another extension of this paper could add more insight by accounting for a stochastic innovation process, market entry and technology switching cost.

[17] In the renewable energy sector early installations increase the positive externality. The total amount of carbon dioxide in the atmosphere is reduced, because renewable energy sources are substitutes to conventional energy sources that emit $C0_2$.

[18] For example in medium-sized countries, there is a fixed number of places, where wind energy plants can be built.

6 Appendix

Proof of lemma 2

A transversality condition for the state variable solves for $const_1$ in (13). The total cost before t_1 is sunk and can be ignored, $\eta_0(t_1)$ has an influence only on future cost. The area below the learning curve between $y(t_1)$ and $y(t)$, for $t \in [t_1, T]$ is defined as

$$\Gamma\left[y(t)\right] = \int_{y(t_1)}^{y(t)} m + ce^{-\lambda y(t_1) - \gamma[v - y(t_1)]}\, dv = m[y(t) - y(t_1)] + \frac{c}{\gamma}[e^{-\lambda y(t_1)} - e^{-\lambda y(t_1) - \gamma[y(t) - y(t_1)]}] \tag{55}$$

The time derivative is $\dot{\Gamma} = \frac{d\Gamma}{dt} = x(t)[m + ce^{-\lambda y(t_1) - \gamma[y(t) - y(t_1)]}]$, thus the time dependent area under the learning curve or the total cost of the second phase is

$$\int_{t_1}^{T} \dot{\Gamma} dt = [y(T) - y(t_1)]m + \frac{c}{\gamma}[e^{-\lambda y(t_1)} - e^{-\lambda y(t_1) - \gamma[y(T) - y(t_1)]}]$$

The usual methods of the principle of variations are used. At first the optimal path of the production flow is displaced for the cost that occurs after t_1; $x(p_1(t), t) \to x(p_1(t), t) + \delta\phi(t)$, $-\frac{\lambda}{\gamma}ce^{-\lambda y(t_1)} + \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1) - \gamma y(T)}$

$$\Gamma\left[y(t)\right] = y(T)m - y(t_1)m - \frac{1}{\gamma}ce^{(\gamma-\lambda)y(t_1) - \gamma y(T)} + \frac{1}{\gamma}ce^{-\lambda y(t_1)}$$

$$\Gamma\left[y(t)\right]^{\delta} = [\int_{t_0}^{T} [x(t) + \delta\phi(t)]dt]m - [\int_{t_0}^{t_1} [x(t) + \delta\phi(t)]dt]m$$

$$+\frac{1}{\gamma}ce^{-\lambda\int_{t_0}^{t_1}[x(t)+\delta\phi(t)]dt} - \frac{1}{\gamma}ce^{(\gamma-\lambda)\int_{t_0}^{t_1}[x(t)+\delta\phi(t)]dt - \gamma\int_{t_0}^{T}[x(t)+\delta\phi(t)]dt}$$

In a second step the displaced total cost after t_1 (called $\Gamma\left[y(t)\right]^{\delta}$) is differentiated with respect to δ. The derivative is evaluated at $\delta = 0$, employing the standard calculus of variations approach,

$$\frac{\partial\Gamma\left[y(t)\right]^{\delta}}{\partial\delta}|_{(\delta=0)} = \left\{-\frac{\lambda}{\gamma}ce^{-\lambda y(t_1)} + \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1) - \gamma y(T)}\right\}\int_{t_0}^{t_1}\phi(t)dt \tag{56}$$

$$+\{m + ce^{(\gamma-\lambda)y(t_1) - \gamma y(T)}\}\int_{t_1}^{T}\phi(t)dt$$

where (56) separates the terms multiplied by $\int_{t_0}^{t_1}\phi(t)dt$ and $\int_{t_1}^{T}\phi(t)dt$ respectively. For any $t > t_1$, the marginal cost at T are collected. This result is similar to Spence's (1981) learning model. Given a production plan, he shows that when a firm extends its production by one unit at any time, then the incremental cost is equal to marginal cost at the end of the planning horizon. For $t_0 \le t \le t_1$ we collect another term; $-\frac{\lambda}{\gamma}ce^{-\lambda y(t_1)} + \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1) - \gamma y(T)} < 0$. For $t = t_1$ the incremental cost is thus the marginal cost of the second phase at T plus $\left[-\frac{\lambda}{\gamma}ce^{-\lambda y(t_1)} + \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)}\right]$. The value of delaying the innovation by one production unit is obtained when this term is multiplied by -1. Thus the first phase's costate at the innovation date is $\eta_0(t_1) = \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)} - \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)}$.

Proof of proposition 3

When a firm produces an additional quantity ϵ_0 *before* the innovation occurs $(t_0 \leq t \leq t_1)$ then the pricing rule (24) holds. The incremental cost for the first time phase is computed next. The total cost is equal to

$$y(T)m + \frac{c}{\lambda}[1 - e^{-\lambda y(t_1)}] + \frac{c}{\gamma}[e^{-\lambda y(t_1)} - e^{-\lambda[y(t_1)]-\gamma[y(T)-y(t_1)]}] \tag{57}$$

When the firm produces an additional unit ϵ_0 before the innovation takes place, then total cost increases to

$$(y(T)+\epsilon_0)m + \frac{c}{\lambda}[1 - e^{-\lambda[y(t_1)+\epsilon_0]}] + \frac{c}{\gamma}[e^{-\lambda[y(t_1)+\epsilon_0]} - e^{-\lambda[y(t_1)+\epsilon_0]-\gamma[y(T)-y(t_1)]}] \tag{58}$$

The incremental cost during the first phase is denoted by IC_0

$$IC_0 = \frac{(58)-(57)}{\epsilon_0} = c\{-\frac{e^{-\lambda[y(t_1)+\epsilon_0]}}{\lambda\epsilon_0} - e^{-\lambda y(t_1)}] + \frac{e^{-\lambda[y(t_1)+\epsilon_0]} - e^{-\lambda y(t_1)}}{\gamma\epsilon_0}.$$

$$-\frac{e^{-\lambda[y(t_1)+\epsilon_0]-\gamma[y(T)-y(t_1)]} - e^{-\lambda[y(t_1)]-\gamma[y(T)-y(t_1)]}}{\gamma\epsilon_0}\} + m \tag{59}$$

It needs to be shown that (24) is equivalent to setting the marginal revenue equal to the incremental cost (59). $ce^{-\lambda y(t_1)}$ is equivalent to $\lim\limits_{\epsilon_0->0} -\frac{c}{\lambda\epsilon_0}[e^{-\lambda[y(t_1)+\epsilon_0]} - e^{-\lambda y(t_1)}]$ and $-\frac{\lambda}{\gamma}e^{-\lambda y(t_1)}$ to $\lim\limits_{\epsilon_0->0} \frac{e^{-\lambda[y(t_1)+dx]}-e^{-\lambda y(t_1)}}{\gamma\epsilon_0}$. In order to show the equivalence of the third term, it has to be transformed: $\lim\limits_{\epsilon_0->0} -\frac{1}{\gamma\epsilon_0}[e^{-\lambda[y(t_1)+\epsilon_0]-\gamma[y(T)-y(t_1)]} - e^{-\lambda[y(t_1)]-\gamma[y(T)-y(t_1)]}] = -\frac{1}{\gamma\epsilon_0}[e^{(\gamma-\lambda)y(t_1)-\gamma y(T)-\lambda\epsilon_0} - e^{-\lambda[y(t_1)]-\gamma[y(T)-y(t_1)]}]$. It follows that $\lim\limits_{\epsilon_0->0} -\frac{1}{\gamma\epsilon_0}[e^{(\gamma-\lambda)y(t_1)-\gamma y(T)-\lambda\epsilon_0} - e^{-\lambda[y(t_1)]-\gamma[y(T)-y(t_1)]}] = \frac{\lambda}{\gamma}e^{(\gamma-\lambda)y(t_1)-\gamma y(T)}$, where the following rule was applied: $\lim\limits_{\epsilon_0->0} \frac{e^{a-\lambda\epsilon_0}-e^{a}}{\epsilon_0} = e^a \lim\limits_{\epsilon_0->0} \frac{e^{-\lambda\epsilon}-1}{\epsilon} = e^a\lambda$, for any constant $a \in \mathbb{R}$. This completes the first part of the proof. During the first time phase, the monopolist behaves optimally, when it sets marginal revenue equal to incremental cost at each instant of time. Next, the second time phase $\nabla t_1 \leq t \leq T$ is analyzed. (25) is the optimality condition during this phase. When the firm produces an additional unit ϵ_1 *after* the innovation has occurred, total cost (57) increases to

$$(y(T)+\epsilon_1)m + \frac{c}{\lambda}[1 - e^{-\lambda y(t_1)}] + \frac{c}{\gamma}[e^{-\lambda y(t_1)} - e^{-\lambda y(t_1)-\gamma[y(T)+\epsilon_1-y(t_1)]}] \tag{60}$$

IC_1 is the incremental cost that occurs through an additional ϵ_1 after t_1

$$IC_1 = \frac{(60)-(57)}{\epsilon_1} = \frac{c_0}{\gamma\epsilon_1}[e^{-\lambda y(t_1)-\gamma[y(T)-y(t_1)]} - e^{-\lambda y(t_1)-\gamma[y(T)+\epsilon_1-y(t_1)]}] \tag{61}$$

(25) is equivalent to (61), because $ce^{(\gamma-\lambda)y(t_1)-\gamma[y(T)-y(t_1)]} = \lim_{\epsilon_1 -> 0} \frac{c}{\gamma\epsilon_1}[e^{-\lambda y(t_1)-\gamma[y(T)-y(t_1)]} - e^{-\lambda y(t_1)-\gamma[y(T)+\epsilon_1-y(t_1)]}]$.

Proof of proposition 4

Pricing during the first phase ($t \leq t_1$); when $r \neq 0$

This section reconstructs equation (24) in the presence of discounting with $t_0 = 0$. The costate that contains a discount rate is denoted by $\psi_0(t) = e^{-rt}$ $\eta_0(t)$. The transversality condition for $\psi_0(t_1)$ has to be derived. It is shown that (12) holds, when $\eta_0(t)$ is substituted for $\psi_0(t)$

$$\frac{\partial H}{\partial p_0(t)} = x_0(p_0(t),t)+p_0(t)\frac{\partial x_0(p_0(t),t)}{\partial p_0(t)} - \frac{\partial x_0(p_0(t),t)}{\partial p_0(t)}(m+ce^{-\lambda y(t)})+\psi_0(t)\frac{\partial x_0(p_0(t),t)}{\partial p_0(t)} \doteq 0$$

$$\Leftrightarrow p_0(1-\frac{1}{\varepsilon(t)}) = m + ce^{-\lambda y(t)} - \psi_0(t) \tag{62}$$

The second first order condition (13), becomes

$$\dot{\psi}_0(t) - r\psi_0(t) = -\frac{\partial H}{\partial y_0(t)}$$

multiplying this equation by e^{-rt} and transforming yields

$$\Leftrightarrow e^{-rt}\dot{\psi}_0(t) - re^{-rt}\psi_0(t) = -e^{-rt}\frac{\partial H}{\partial y_0(t)}$$

$$\Leftrightarrow e^{-rt}\psi_0(t) + const_3 = \int \left[e^{-rt}(-\lambda)ce^{-\lambda y_0(t)}x_0(t)\right] dt$$

$$\Leftrightarrow e^{-rt}\psi_0(t) + const_3 = e^{-rt}ce^{-\lambda y_0(t)} + r\int e^{-rt}ce^{-\lambda y_0(t)}dt$$

$$\Leftrightarrow \psi_0(t) = ce^{-\lambda y_0(t)} + re^{rt}\int e^{-rt}ce^{-\lambda y_0(t)}dt - e^{rt}const_3 \tag{63}$$

In order to find an expression of the costate at the innovation date $\psi_0(t_1)$, the total cost that occurs after the innovation is expressed by V.

$$V = \int_{t_1}^{T} \left\{x_1(p_1(t)(m + ce^{(\gamma-\lambda)y(t_1)-\gamma y(t)})e^{-rt}\right\} dt$$

Integration by parts yields:

$$V = \int_{t_1}^{T} re^{-rt}\left\{my_1(t) - \frac{1}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(t)}\right\} dt$$

$$+e^{-rt}\left\{my(T) - \frac{1}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - my(t_1) + \frac{1}{\gamma}ce^{-\lambda y(t_1)}\right\}$$

Applying the usual methods of the principle of variations; the optimal path of the production flow $x(p_1(t),t) \rightarrow x(p_1(t),t) + \delta\phi(t)$ in V is displaced, before its derivative is evaluated at $\delta = 0$ according to the standard calculus of variations approach.

$$V = \int_{t_1}^{T} re^{-rt} \left\{ m \int_{t_0}^{t} [x_1(t) + \delta\phi(t)]\, dt - \frac{1}{\gamma} ce^{(\gamma-\lambda)\int_{t_0}^{t_1}[x_1(t)+\delta\phi(t)]dt - \gamma \int_{t_0}^{t}[x_1(t)+\delta\phi(t)]dt} \right\} dt$$

$$+e^{-rt}\{m \int_{t_0}^{T} [x_1(t) + \delta\phi(t)]\, dt - \frac{1}{\gamma} ce^{(\gamma-\lambda)\int_{t_0}^{t_1}[x_1(t)+\delta\phi(t)]dt - \gamma \int_{t_0}^{T}[x_1(t)+\delta\phi(t)]dt}$$

$$-m \int_{t_0}^{t_1} [x_1(t) + \delta\phi(t)]\, dt + \frac{1}{\gamma} ce^{-\lambda \int_{t_0}^{t_1}[x_1(t)+\delta\phi(t)]dt}\}$$

$$\frac{\partial V}{\partial \delta}]_{\delta=0} = \int_{t_1}^{T} re^{-rt} \left\{ m \int_{t_0}^{t} \phi(t)dt - \frac{\gamma-\lambda}{\gamma} ce^{(\gamma-\lambda)y(t_1)-y(t)} \int_{t_0}^{t_1} \phi(t)dt + ce^{(\gamma-\lambda)y(t_1)-y(t)} \int_{t_0}^{t} \phi(t)dt \right\} dt$$

$$+e^{-rt}\{m \int_{t_0}^{T} \phi(t)dt - \frac{\gamma-\lambda}{\gamma} ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} \int_{t_0}^{t_1} \phi(t)dt + ce^{(\gamma-\lambda)y(t_1)-y(T)} \int_{t_0}^{T} \phi(t)dt$$

$$-m \int_{t_0}^{t_1} \phi(t)dt - \frac{\lambda}{\gamma} ce^{-\lambda y(t_1)} \int_{t_0}^{t_1} \phi(t)dt\}$$

$$= \int_{t_1}^{T} re^{-rt} \left\{ \left[m + ce^{(\gamma-\lambda)y(t_1)-y(t)} \right] \int_{t_0}^{t} \phi(\tau)d\tau - \frac{\gamma-\lambda}{\gamma} ce^{(\gamma-\lambda)y(t_1)-y(t)} \int_{t_0}^{t_1} \phi(t)dt \right\} dt$$

$$+e^{-rt}\{ \int_{t_0}^{t_1} \left[\frac{\lambda}{\gamma} ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - \frac{\lambda}{\gamma} ce^{-\lambda y(t_1)} \right] \phi(t)dt + \int_{t_1}^{T} \left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)} \right] \phi(t)dt\} \quad (64)$$

The term $\int_{t_1}^{T} re^{-rt} \left\{ \frac{\gamma-\lambda}{\gamma} ce^{(\gamma-\lambda)y(t_1)-y(t)} \int_{t_0}^{t_1} \phi(t)dt \right\} dt$ can be ignored. It originates from displacing the cumulative quantity before t_1 after the innovation has already taken place. As the firm cannot change $y(t_1)$ after t_1, it cannot effect the cost during the first time phase at t_1. It just has an effect on the second phase's cost. One can also examine this term itself and recognize that for any t it is zero, because when $t < t_1$ then the first integral becomes zero, when $t > t_1$, then the second integral is zero. Hence (64) becomes

$$\int_{t_1}^{T} re^{-rt} \left\{ \left[m + ce^{(\gamma-\lambda)y(t_1)-y(t)} \right] \int_{t_0}^{t} \phi(\tau)d\tau \right\} dt$$

$$+e^{-rt}\{ \int_{t_0}^{t_1} \left[\frac{\lambda}{\gamma} ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - \frac{\lambda}{\gamma} ce^{-\lambda y(t_1)} \right] \phi(t)dt + \int_{t_1}^{T} \left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)} \right] \phi(t)dt\} \quad (65)$$

One can replace the variables of the first term above. The range of τ is $[t_0, t]$ and that of t is: $[t_1, T] \;\rightarrow t_0 \leq \tau \leq t \leq T$. After replacing the variables $t_0 \leq t \leq \tau \leq T$, the range is $[t, T]$ for τ, when $t > t_1$ and it is $[t_1, T]$ when

$t \leq t_1$. Before the swap, t was larger than t_1 hence τ is larger than t_1 after the swap. When t is smaller than t_1, then τ's lower limit is t_1. When t is larger than t_1, then τ's lower limit is t. The new range of t is $[t_0, T]$. Figure 6 illustrates the range before and after replacing the variables t and τ.
(65) can be transformed to

$$\int_{t_0}^{t_1} re^{-rt}\left\{\left[m + ce^{(\gamma-\lambda)y(t_1)-y(t)}\right]\int_{t_0}^{t}\phi(\tau)d\tau\right\}dt+$$
$$\int_{t_1}^{T} re^{-rt}\left\{\left[m + ce^{(\gamma-\lambda)y(t_1)-y(t)}\right]\int_{t_0}^{t}\phi(\tau)d\tau\right\}dt$$
$$+e^{-rt}\{\int_{t_0}^{t_1}\left[\frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)}\right]\phi(t)dt+\int_{t_1}^{T}\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]\phi(t)dt\}$$

The variables are replaced next.

$$= \int_{t_0}^{t_1}\left\{\int_{t_1}^{T} r\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]e^{-r\tau}d\tau\right\}\phi(t)dt+\int_{t_1}^{T}\left\{\int_{t}^{T} r\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]e^{-r\tau}d\tau\right\}$$
$$\phi(t)dt+e^{-rt}\{\int_{t_0}^{t_1}\left[\frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)}\right]\phi(t)dt+\int_{t_1}^{T}\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]\phi(t)dt\}$$
$$= \int_{t_0}^{t_1}\left\{\int_{t_1}^{T} r\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]e^{-r\tau}d\tau\right\}\phi(t)dt$$
$$+e^{-rt}\{\int_{t_0}^{t_1}\left[\frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)}\right]\phi(t)dt$$
$$+\int_{t_1}^{T}\left\{\int_{t}^{T} r\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]e^{-r\tau}d\tau\right\}\phi(t)dt+\int_{t_1}^{T} e^{-rt}\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]\phi(t)dt\}$$

$$= \int_{t_0}^{t_1}\left\{e^{-rt}\left[\frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)}\right] + \int_{t_1}^{T} r\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]e^{-r\tau}d\tau\right\}\phi(t)dt$$
$$+\int_{t_1}^{T}\left\{e^{-rt}\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right] + \int_{t}^{T} r\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]e^{-r\tau}d\tau\right\}\phi(t)dt\} \tag{66}$$

The same reasoning as in (56) applies here hence the costate of the first phase with a non-zero discount rate at t_1 equals,

$$\rightarrow \psi_0(t_1) = e^{-rt_1}\left[\frac{\lambda}{\gamma}ce^{-\lambda y(t_1)} - \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)}\right] - \int_{t_1}^{T} r\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]e^{-r\tau}d\tau \tag{67}$$

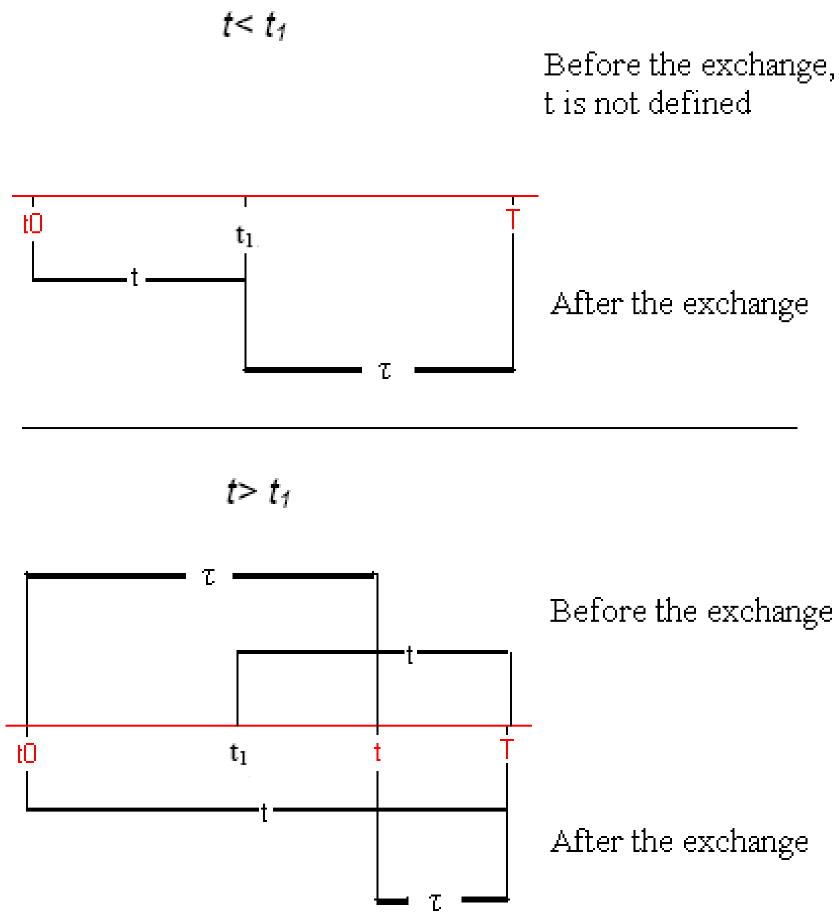

Figure 6: Exchanging variables

The constant $const_3$ is found in the following.

$$\Leftrightarrow \psi_0(t) = ce^{-\lambda y_0(t)} + re^{rt}\int e^{-rt}ce^{-\lambda y_0(t)}dt - e^{rt}const_3$$

$$\rightarrow \psi_0(t_1) = ce^{-\lambda y_0(t_1)} + \left[re^{rt}\int e^{-rt}ce^{-\lambda y_0(t)}dt\right]_{t=t_1} - e^{rt_1}const_3$$

$$\rightarrow ce^{-\lambda y_0(t_1)} + \left[re^{rt}\int e^{-rt}ce^{-\lambda y_0(t)}dt\right]_{t=t_1} - e^{rt_1}const_3 =$$

$$e^{-rt_1}\left[\frac{\lambda}{\gamma}ce^{-\lambda y(t_1)} - \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)}\right] - \int_{t_1}^{T} r\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]e^{-r\tau}d\tau$$

$$-\int_{t_1}^{T} r\left[m + ce^{(\gamma-\lambda)y(t_1)-y(T)}\right]e^{-r\tau}d\tau$$

$$\Leftrightarrow e^{rt_1}const_3 = \left[\frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)}\right]e^{-rt_1} + \int_{t_1}^{T} r[m+ce^{(\gamma-\lambda)y(t_1)-\gamma y(\tau)}]e^{-r\tau}d\tau$$

$$+ce^{-\lambda y_0(t_1)} + \left[re^{rt}\int e^{-rt}ce^{-\lambda y_0(t)}dt\right]_{t=t_1}$$

It is substituted into the costate function (67), which is

$$\rightarrow \psi_0(t) = ce^{-\lambda y_0(t)} + re^{rt}\int e^{-rt}ce^{-\lambda y_0(t)}dt$$

$$-\left[\frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)}\right]e^{-rt_1} - \int_{t_1}^{T} r[m+ce^{(\gamma-\lambda)y(t_1)-\gamma y(\tau)}]e^{-r\tau}d\tau - ce^{-\lambda y_0(t_1)} \tag{68}$$

Finally one can substitute the costate in equation (68) to find an expression, how the monopolist sets the price during the first time phase.

$$p_0(1-\frac{1}{\varepsilon(t)}) = m + ce^{-\lambda y_0(t)} - ce^{-\lambda y_0(t)} - re^{rt}\int e^{-rt}ce^{-\lambda y_0(t)}dt$$

$$+\left[\frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)}\right]e^{-rt_1} + \int_{t_1}^{T} r[m+ce^{(\gamma-\lambda)y(t_1)-\gamma y(\tau)}]e^{-r\tau}d\tau + ce^{-\lambda y_0(t_1)}$$

$$\left[re^{rt}\int e^{-rt}ce^{-\lambda y_0(t)}dt\right]_{t=t_1}$$

$$\Leftrightarrow MR_0(t) = m + \left[\frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} - \frac{\lambda}{\gamma}ce^{-\lambda y(t_1)}\right]e^{-rt_1} + ce^{-\lambda y_0(t_1)}$$

$$+\left[re^{rt}\int e^{-rt}ce^{-\lambda y_0(t)}dt\right]_{t=t_1} - re^{rt}\int e^{-rt}ce^{-\lambda y_0(t)}dt + r\int_{t_1}^{T}[m+ce^{(\gamma-\lambda)y(t_1)-\gamma y(\tau)}]e^{-r\tau}d\tau$$

$$\Leftrightarrow MR_0(t) = m + \frac{\lambda}{\gamma}ce^{(\gamma-\lambda)y(t_1)-\gamma y(T)} + \frac{\gamma-\lambda}{\gamma}ce^{-\lambda y(t_1)}$$

$$+r\int_{t}^{t_1}\left[m+ce^{-\lambda y_0(\tau)}\right]e^{-r(\tau-t)}d\tau + r\int_{t_1}^{T}[m+ce^{(\gamma-\lambda)y(t_1)-\gamma y(\tau)}]e^{-r(\tau-t)}d\tau \tag{69}$$

With a positive discount rate, there are two additional terms in the optimality condition. These reflect the fact that the return to experience or learning, which is expressed by a lower future cost, matters less, due to the presence of discounting. The optimal plan is to increase p_0 and thus present profits at the expense of future profits. Note that (69) reduces to (24), when $r = 0$.

Pricing during the second phase ($t > t_1$); when $r \neq 0$

(25) and (17) can be rewritten to account for discounting

$$MR_1(t) = m + ce^{(\gamma-\lambda)y(t_1)-\gamma y(t)} - \psi_1(t) \tag{70}$$

$$\psi_1(t) = ce^{(\gamma-\lambda)y(t_1)-\gamma y_1(t)} + re^{rt}\int e^{-rt}e^{(\gamma-\lambda)y(t_1)-\gamma y(t)}dt - e^{-rt}const_4 \tag{71}$$

Extending production at $t = T$ does not increase the profit, if the firm behaves optimally. Hence $\psi_1(T) = 0$.

$$\rightarrow \psi_1(T) = ce^{(\gamma-\lambda)y(t_1)-\gamma y_1(T)} + \left[re^{rt}\int e^{-rt}e^{(\gamma-\lambda)y(t_1)-\gamma y(t)}dt\right]_{t=T} - e^{-rT}const_4 = 0$$

$$\Leftrightarrow ce^{(\gamma-\lambda)y(t_1)-\gamma y_1(T)} + \left[re^{rt}\int e^{-rt}e^{(\gamma-\lambda)y(t_1)-\gamma y(t)}dt\right]_{t=T} = e^{-rT}const_4$$

$$\rightarrow \psi_1(t) = ce^{(\gamma-\lambda)y(t_1)-\gamma y_1(t)} + re^{rt}\int e^{-rt}e^{(\gamma-\lambda)y(t_1)-\gamma y(t)}dt - ce^{(\gamma-\lambda)y(t_1)-\gamma y_1(T)}$$

$$-\left[re^{rt}\int e^{-rt}e^{(\gamma-\lambda)y(t_1)-\gamma y(t)}dt\right]_{t=T}$$

$$MR_1(t) = m + ce^{(\gamma-\lambda)y(t_1)-\gamma y_1(T)} - re^{rt}\int e^{-rt}e^{(\gamma-\lambda)y(t_1)-\gamma y(t)}dt + \left[re^{rt}\int e^{-rt}e^{(\gamma-\lambda)y(t_1)-\gamma y(t)}dt\right]_{t=T}$$

$$\Leftrightarrow MR_1(t) = m + ce^{(\gamma-\lambda)y(t_1)-\gamma y_1(T)} - re^{rt}\int e^{-rt}e^{(\gamma-\lambda)y(t_1)-\gamma y(t)}dt + \left[re^{rt}\int e^{-rt}e^{(\gamma-\lambda)y(t_1)-\gamma y(t)}dt\right]_{t=T}$$

During the second phase, the monopolist sets p_1 such that

$$\Leftrightarrow MR_1(t) = m + ce^{(\gamma-\lambda)y(t_1)-\gamma y_1(T)} + r\int_{t}^{T}\left[m+ce^{(\gamma-\lambda)y(t_1)-\gamma y_1(\tau)}\right]e^{-r(\tau-t)}d\tau \tag{72}$$

With a positive discount rate, there is an additional term. Again, the optimality condition reflects the fact that the return to experience is in the future, which is discounted, thus p_1 rises. (72) reduces to (25), when $r = 0$.

Proof of proposition 6

For a fix total production quantity $\overline{y(T)}$, it is shown analytically and numerically that a subsidy on innovation cost, induces the innovation to occur earlier and thus p_0 to increase. The welfare loss could be encountered by a price ceiling during this phase. The price ceiling would have a counter effect, when the date of innovation is delayed by the introduction of a price constraint. In this case, consumers would pay the higher price p_0 for a longer period of time. The answer is a straightforward extension of the preceding section and is based on a method that is used in Rees (1986). p_0 is constant, therefore either the price constraint $\overline{p} - p_0 \geq 0$, where $\overline{p}$ is the price ceiling binds over the entire interval $[t_0, t_1]$ or it does not bind at all. When it does not bind, it has no effect. Assume it does bind, and $p_0^* > \overline{p}$, where p_0^* is the optimal price set by the monopolist; in this case the monopolist looses some of its profit due to the price cap. This loss is denoted by $R[x_0(\overline{p})]$, with $R_{x_0} < 0$, and $R > 0$ if $x < x(\overline{p_0})$. The per period quantity x_0 on which R depends upon is considered as a function of $\overline{p}$. Whenever $p_0 \leq \overline{p}$, then $x > x(\overline{p_0})$ and $R \equiv 0$. The new objective is $\underset{x_0, x_1, t_1}{Max}\ MP' = \displaystyle\int_{t_0}^{t_1} \left[\pi^0(t) - R(x(\overline{p}))\right] dt + \int_{t_1}^{T} \pi^1(t)dt - a(t_1)$ subject to the constraints (5)-(9). The problem can be divided into two segments. The maximand over the first phase $[t_0, t_1]$ is $V_0^* \equiv Max \int_{t_0}^{t_1} \left[\pi^0(t) - R(x_0(\overline{p_0}))\right] dt - a(t_1)$, that of the second phase is $V_1^* \equiv Max \displaystyle\int_{t_1}^{T} \pi^1(t)dt$. t_1^* maximizes the sum $V_0^* + V_1^*$, and satisfies $\partial\left(V_0^* + V_1^*\right)/t_1 \overset{\circ}{=} 0$, from which the optimality condition $a'(t_1^*) + R(x(\overline{p_0})) = \pi^0(t_1^*) + \eta_0(t_1^*)x[p_0(t_1^*), t_1^*] - \pi^1(t_1^*) - \eta_1(t_1^*)x[p_1(t_1^*), t_1^*]$ can be derived. Aside from R the optimality condition is the same as (39). The introduction of a price constraint decreases t_1^*. The low price p_1 is charged for a longer time span, because the length of phase 2 increases. In addition, an earlier innovation date causes p_1 to decrease further.

Summing up one can say that the introduction of a price constraint along with an innovation subsidy increases consumer rents during the entire planning horizon. Producers receive lower revenues, when $\overline{y(T)}$ is fixed, because prices do not increase, however, their innovation cost declines, due to the subsidy. Their production cost decreases, because producers learn with a larger learning parameter sooner.

7 References

Brueckner, Jan and Raymon, Neil (1983), "Optimal Production with Learning by Doing", *Journal of Economic Dynamics and Control. Vol. 8, 127-135*

Dena (2005), "Energiewirtschaftliche Planung für die Netzintegration von Windenergie in Deutschland an Land und Offshore bis zum Jahr 2020", Internet address: http://www.offshore-wind.de, accessed on 23.03.2008

Irwin, Douglas and Klenow, Peter (1994), "Learning-by-Doing Spillovers in the Semiconductor Industry," *Journal of Political Economy, Vol.102, 1200-27*

Iset (2005), "Effizienz der Windenergie, Auszug aus: Windenergie Report Deutschland 2005", Internet address: http://www.wind-energie.de, accessed on 10.12.2009

Parente, Stephen (1994), "Technology Adoption, Learning-by Doing and Economic Growth", *Journal of Economic Theory, Vol. 63, 346-369*

Spence, Michael (1977), "Entry, investment and oligopolistic pricing" *Bell Journal of Economics, Vol. 8, 534-44*

Spence, Michael (1981), "The Learning Curve and Competition", *Bell Journal of Economics, Vol. 12, 49-70*

Takayama, Akira (1974), "Mathematical Economics", *Hinsdale Illinois, The Dryden Press*
Young, Alwyn (1993), "Invention and Bounded Learning by Doing", *The Journal of Political Economy, Vol. 101, 443-472*